ENGLISH TRANSLATION OF INTERNATIONAL PATENT APPLICATION PCT/FR97/00214

filed on February 3, 1997

in the name of SANOFI

SANOFI

PURIFIED SR-p70 PROTEIN

Abstract

The invention relates to new nucleic acid sequences of the family of tumour-suppressing genes related to the gene for the p53 protein, and to the corresponding protein sequences.

Th invention relates to new nucleic acid sequences of the family of tumour-suppressing genes r lated to the gene for the p53 protein, and to the corresponding protein sequences.

The invention also relates to the prophylactic, therapeutic and diagnostic applications of thes sequences, in particular in the field of pathologies linked to the phenomena of apoptosis or of cell transformation.

Tumour-suppressing genes perform a key role in protection against the phenomena of carcinogenesis, and any modification capable of bringing about the loss of one of these genes, its inactivation or its dysfunction may have oncogenic character, thereby creating favourable conditions for the development of a malignant tumour.

The authors of the present invention have identified transcription products of a new gene, as well as the corresponding proteins. This gene, SR-p70, is related to the p53 tumour-suppressing gene, the antitumour activity of which is linked to its transcription factor activity, and more specifically to the controls exerted on the activity of the Bax and Bcl-2 genes which are instrumental in the mechanisms of cell death.

Hence the present invention relates to purified SR-p70 proteins, or biologically active fragments of the latter.

The invention also relates to isolated nucleic acid sequences coding for the said proteins or their biologically active fragments, and to specific oligonucleotides obtained from these sequences.

It relates, in addition, to the cloning and/or expression vectors containing at least one of the nucleotide sequences defined above, and the host cells transfected by these cloning and/or expression vectors under conditions permitting the replication and/or expression of one of th said nucl otide sequ nc s.

The methods of production of recombinant SR-p70 proteins or their biologically activ fragments by the transfect d host cells also f rm part of the invention.

The invention also comprises antibodies or antibody derivatives specific for the proteins defined above.

It relates, in addition, to methods of detection of cancers, either by measuring the accumulation of SR-p70 proteins in the tumours according to immunohistochemical techniques, or by demonstrating autoantibodies directed against these proteins in patients' serum.

The invention also relates to any inhibitor or activator of SR-p70 activity, for example of protein-protein interaction, involving SR-p70.

It also relates to antisense oligonucleotide sequences specific for the above nucleic acid sequences, capable of modulating *in vivo* the expression of the SR-p70 gene.

Lastly, the invention comprises a method of gene therapy, in which vectors such as, for example, inactivated viral vectors capable of transferring coding sequences for a protein according to the invention are injected into cells deficient for this protein, for purposes of regulating the phenomena of apoptosis or of reversion of transformation.

A subject of the present invention is a purified polypeptide comprising an amino acid sequence selected from:

a) the sequence SEQ ID No. 2;
b) the sequence SEQ ID No. 4;
c) the sequence SEQ ID No. 6;
d) the sequence SEQ ID No. 8;
e) the sequence SEQ ID No. 10;
f) the sequence SEQ ID No. 13;
g) the sequence SEQ ID No. 15;
h) the sequence SEQ ID No. 17;
i) the sequence SEQ ID No. 19;
j) any biologically active sequence derived from SEQ ID No. 2, SEQ ID No. 4, SEQ ID No. 6, SEQ ID No. 8, SEQ ID No. 10, SEQ ID No. 13, SEQ ID No. 15, SEQ ID No. 17 or SEQ ID No. 19.

In the description of the invention, the following definitions are used:

- SR-p70 protein: a polypeptide comprising an amino acid sequence selected from the sequences SEQ ID No. 2, SEQ ID No. 4, SEQ ID No. 6, SEQ ID No. 8, SEQ ID No. 10, SEQ ID No. 13, SEQ ID No. 15, SEQ ID No. 17 or SEQ ID No. 19, or any biologically active fragment or derivative of this polypeptide;

- derivative: any variant polypeptide of the polypeptide of sequence SEQ ID No. 2, SEQ ID No. 4, SEQ ID No. 6, SEQ ID No. 8, SEQ ID No. 10, SEQ ID No. 13, SEQ ID No. 15, SEQ ID No. 17 or SEQ ID No. 19, or any molecule resulting from a modification of a genetic and/or chemical nature of the sequence SEQ ID No. 2, SEQ ID No. 4, SEQ ID No. 6, SEQ ID No. 8, SEQ ID No. 10, SEQ ID No. 13, SEQ ID No. 15, SEQ ID No. 17 or SEQ ID No. 19, that is to say obtained by mutation, deletion, addition, substitution and/or chemical modification of a single amino acid or of a limited number of amino acids, as well as any isoform sequence, that is to say sequence identical to the sequence SEQ ID No. 2, SEQ ID No. 4, SEQ ID No. 6, SEQ ID No. 8, SEQ ID No. 10, SEQ ID No. 13, SEQ ID No. 15, SEQ ID No. 17 or SEQ ID No. 19, or to one of its fragments or modified sequences, containing one or more amino acids in the form of the D enantiomer, the said variant, modified or isoform sequences having retained at least one of the properties that make them biologically active;

- biologically active: capable of binding to DNA and/or of exerting transcription factor activity and/or of participating in the control of the cell cycle, of differentiation and of apoptosis and/or capable of being recognized by the antibodies specific for the polypeptide of sequence SEQ ID No. 2, SEQ ID No. 4, SEQ ID No. 6, SEQ ID No. 8, SEQ ID N . 10, SEQ ID No. 13, SEQ ID No. 15, SEQ ID No. 17 or SEQ ID No. 19, and/or capable of inducing antibodies which recognize this polypeptide.

The manufactur of derivatives may hav different objectives, including esp cially that of increasing the

affinity of th polypeptide for DNA or its transcription factor activity, and that of improving its levels of production, of increasing its resistance to proteases, of modifying its biological activities or of endowing it with new pharmaceutical and/or biological properties.

Among the polypeptides of the invention, the polypeptide of human origin comprising the sequence SEQ ID No. 6, SEQ ID No. 13, SEQ ID No. 15, SEQ ID No.17 or SEQ ID No. 19 is preferred. The polypeptide of 636 amino acids corresponding to the sequence SEQ ID No. 6 is more than 97% identical to the polypeptide of sequence SEQ ID No. 2.

The polypeptide of sequence SEQ ID No. 2 and that of sequence SEQ ID No. 4 are two expression products of the same gene, and the same applies to the sequences SEQ ID No. 8 and SEQ ID No. 10 and to the sequences SEQ ID No. 6, SEQ ID No. 13, SEQ ID No. 15, SEQ ID No. 17 or SEQ ID No. 19.

As will be explained in the examples, the polypeptide of sequence SEQ ID No. 4 corresponds to a premature termination of the peptide of sequence SEQ ID No. 2, linked to an alternative splicing of the longer transcript (messenger RNA), coding for the polypeptide of SEQ ID No. 2, of the corresponding gene. Similarly, in humans, the polypeptides corresponding to the sequences SEQ ID No. 6, SEQ ID No. 13, SEQ ID No. 15, SEQ ID No. 17 and SEQ ID No. 19, diverge in their composition in respect of the N- and/or C-terminal portions, this being the outcome of alternative splicing of the same primary transcript. The N-terminal peptide sequence of the sequence SEQ ID No. 10 is deleted, this being linked to an alternative splicing of its coding transcript.

Advantageously, the invention relates to a polypeptide corresponding to the DNA binding domain of one of the above polypeptides.

This domain corresponds to the sequence lying betwe n residue 110 and residue 310 for the sequenc s SEQ ID No. 2 or 6, and b tween residue 60 and residue 260 for the sequence SEQ ID No. 8.

A subject of the present invention is also nucleic acid sequences coding for a SR-p70 protein or biologically active fragments or derivatives of the latter.

More preferably, a subject of the invention is an isolated nucleic acid sequence selected from:

a) the sequence SEQ ID No. 1;

b) the sequence SEQ ID No. 3;

c) the sequence SEQ ID No. 5;

d) the sequence SEQ ID No. 7;

e) the sequence SEQ ID No. 9;

f) the sequence SEQ ID No. 11;

g) the sequence SEQ ID No. 12;

h) the sequence SEQ ID No. 14;

i) the sequence SEQ ID No. 16;

j) the sequence SEQ ID No. 18;

k) the nucleic acid sequences capable of hybridizing specifically with the sequence SEQ ID No. 1, SEQ ID No. 3, SEQ ID No. 5, SEQ ID No. 7, SEQ ID No. 9, SEQ ID No. 11, SEQ ID No. 12, SEQ ID No. 14, SEQ ID No. 16 or SEQ ID No. 18 or with the sequences complementary to them, or of hybridizing specifically with their proximal sequences;

l) the sequences derived from the sequences a), b), c), d), e), f), g), h), i), j) or k) as a result of the degeneracy of the genetic code.

According to a preferred embodiment, a subject of the invention is nucleotide sequences SEQ ID No. 5, SEQ ID No. 12, SEQ ID No. 14, SEQ ID No. 16 and SEQ ID No. 18, corresponding, respectively, to the cDNAs of the human proteins of the sequences SEQ ID No. 6, SEQ ID No. 13, SEQ ID No. 15, SEQ ID No. 17 and SEQ ID No. 19.

The different nucleotide sequences of the invention may be of artificial origin or otherwise. They can be DNA or RNA sequences obtained by the screening of libraries of sequences by means of probes prepared on the basis of the sequences SEQ ID No. 1, 3, 5, 7, 9, 11, 12, 14, 16 or 18. Such libraries may be prepared by traditional techniques of molecular biology which are

known to a person skilled in the art.

The nucleotide sequences according to the invention may also be prepared by chemical synthesis, or alternatively by mixed methods including the chemical or enzymatic modification of sequences obtained by the screening of libraries.

These nucleotide sequences enable nucleotide probes to be produced which are capable of hybridizing strongly and specifically with a nucleic acid sequence, of a genomic DNA or of a messenger RNA, coding for a polypeptide according to the invention or a biologically active fragment of the latter. Such probes also form part of the invention. They may be used as an *in vitro* diagnostic tool for the detection, by hybridization experiments, of transcripts specific for the polypeptides of the invention in biological samples, or for the demonstration of aberrant syntheses or of genetic abnormalities such as loss of heterozygosity or genetic rearrangement resulting from a polymorphism, from mutations or from a different splicing.

The probes of the invention contain at least 10 nucleotides, and contain at most the whole of the sequence of the SR-p70 gene or of its cDNA contained, for example, in a cosmid.

Among the shortest probes, that is to say of approximately 10 to 20 nucleotides, the appropriat hybridization conditions correspond to the stringent conditions normally used by a person skilled in the art.

The temperature used is preferably between T_m -5°C and T_m -30°C, and as a further preference between T_m -5°C and T_m -10°C, T_m being the melting temperature, the temperature at which 50% of the paired DNA strands separate.

The hybridization is preferably conducted in solutions of high ionic strength, such as, in particular, 6 x SSC solutions.

Advantageously, the hybridization conditions us d are as follows:

- temperature: 42°C,

- hybridization buffer: 6 x SSC, 5 x Denhart's, 0.1% SDS, as describ d in Example III.

Advantageously, these probes are r pr sented by the following oligonucleotides or the sequences complementary to them:

SEQ ID No. 20: GCG AGC TGC CCT CGG AG
SEQ ID No. 21: GGT TCT GCA GGT GAC TCA G
SEQ ID No. 22: GCC ATG CCT GTC TAC AAG
SEQ ID No. 23: ACC AGC TGG TTG ACG GAG
SEQ ID No. 24: GTC AAC CAG CTG GTG GGC CAG
SEQ ID No. 25: GTG GAT CTC GGC CTC C
SEQ ID No. 26: AGG CCG GCG TGG GGA AG
SEQ ID No. 27: CTT GGC GAT CTG GCA GTA G
SEQ ID No. 28: GCG GCC ACG ACC GTG AC
SEQ ID No. 29: GGC AGC TTG GGT CTC TGG
SEQ ID No. 30: CTG TAC GTC GGT GAC CCC
SEQ ID No. 31: TCA GTG GAT CTC GGC CTC
SEQ ID No. 32: AGG GGA CGC AGC GAA ACC
SEQ ID No. 33: CCA TCA GCT CCA GGC TCT C
SEQ ID No. 34: CCA GGA CAG GCG CAG ATG
SEQ ID No. 35: GAT GAG GTG GCT GGC TGG A
SEQ ID No. 36: TGG TCA GGT TCT GCA GGT G
SEQ ID No. 37: CAC CTA CTC CAG GGA TGC
SEQ ID No. 38: AGG AAA ATA GAA GCG TCA GTC
SEQ ID No. 39: CAG GCC CAC TTG CCT GCC
SEQ ID No. 40: CTG TCC CCA AGC TGA TGA G

Preferably, the probes of the invention are labelled prior to their use. To this end, several techniques are within the capacity of a person skilled in the art (fluorescent, radioactive, chemoluminescence, enzyme, and the like, labelling).

The *in vitro* diagnostic methods in which these nucleotide probes are employed are included in the subject of the present invention.

These methods relat , for example, to the detection of abnormal syntheses (.g. accumulation of transcription products) or of genetic abnormalities, such as loss of heterozygosity and genetic rearrangement, and point mutations in the nucl otide sequences of nucleic

acids coding for an SR-p70 protein, according to the definition given above.

Th nucleotide sequenc s of the invention are also useful for the manufacture and use of oligonucleotide primers for sequencing reactions or specific amplification reactions according to the so-called PCR technique or any variant of the latter (ligase chain reaction (LCR), etc).

Preferred primer pairs consist of primers selected from the nucleotide sequences: SEQ ID No. 1: monkey sequence of 2,874 nucleotides, and SEQ ID No. 5: human SR-p70a cDNA, in particular upstream of the ATG translation initiation codon and downstream of the TGA translation stop codon.

Advantageously, these primers are represented by the following pairs:

- pair No. 1:
sense primer: GCG AGC TGC CCT CGG AG (SEQ ID No. 20)
antisense primer: GGT TCT GCA GGT GAC TCA G (SEQ ID No. 21)

- pair No. 2:
sense primer: GCC ATG CCT GTC TAC AAG (SEQ ID No. 22)
antisense primer: ACC AGC TGG TTG ACG GAG (SEQ ID No. 23)

- pair No. 3:
sense primer: GTC AAC CAG CTG GTG GGC CAG (SEQ ID No. 24)
antisense primer: GTG GAT CTC GGC CTC C (SEQ ID No. 25)

- pair No. 4:
sense primer: AGG CCG GCG TGG GGA AG (SEQ ID No. 26)
antisense primer: CTT GGC GAT CTG GCA GTA G (SEQ ID No. 27)

- pair No. 5:
sense primer: GCG GCC ACG ACC GTG A (SEQ ID No. 28)
antisense primer: GGC AGC TTG GGT CTC TGG (SEQ ID No. 29)

- pair No. 6:
sense primer: CTG TAC GTC GGT GAC CCC (SEQ ID No. 30)
antisense primer: TCA GTG GAT CTC GGC CTC (SEQ ID No. 31)

- pair No. 7:
sense primer: AGG GGA CGC AGC GAA ACC (SEQ ID No. 32)
antisense primer: GGC AGC TTG GGT CTC TGG (SEQ ID No. 29)

- pair No. 8: (SEQ ID No. 41)
sense primer: CCCCCCCCCCCCCCN$_1$ (where N equals G, A or T)
antisense primer: CCA TCA GCT CCA GGC TCT C (SEQ ID No. 33)

- pair No. 9: (SEQ ID NO. 41)
sense primer: CCCCCCCCCCCCCCN$_1$ (where N equals G, A or T)
antisense primer: CCA GGA CAG GCG CAG ATG (SEQ ID No. 34)

- pair No. 10: (SEQ ID NO. 42)
sense primer: CCCCCCCCCCCCCCN$_1$ (where N equals G, A or T)
antisense primer: CTT GGC GAT CTG GCA GTA G (SEQ ID No. 27)

- pair No. 11:
sense primer: CAC CTA CTC CAG GGA TGC (SEQ ID No. 37)
antisense primer: AGG AAA ATA GAA GCG TCA GTC (SEQ ID No. 38)

- pair No. 12:
sense primer: CAG GCC CAC TTG CCT GCC (SEQ ID No. 39)
antisense primer: CTG TCC CCA AGC TGA TGA G (SEQ ID No. 40)

These primers correspond to the sequences extending, respectively:

- from nucleotide No. 124 to nucleotide No. 140 on SEQ ID No. 1 and from nucleotide No. 1 to nucleotide No. 17 on SEQ ID No. 5 for SEQ ID No. 20
- from nucleotide No. 2280 to nucleotide No. 2262 on SEQ ID No. 1 and from nucleotide No. 2156 to nucleotide 2138 on SEQ ID No. 5 for SEQ ID No. 21
- from nucleotide No. 684 to nucleotide No. 701 on SEQ ID No. 1 for SEQ ID No. 22
- from nucleotide No. 1447 to nucleotide No. 1430 on SEQ ID No. 1 and from nucleotide 1324 to nucleotide 1307 on SEQ ID No. 5 for SEQ ID No. 23
- from nucleotide 1434 to nucleotide 1454 on SEQ ID No. 1 and from nucleotid 1311 to nucl otide 1331 on SEQ ID N . 5 for SEQ ID No. 24
- from nucl otide 2066 to nucl otide 2051 on SEQ ID No. 1 and from nucleotide 1940 to nucleotide

1925 on SEQ ID No. 5 for SEQ ID No. 25

- from nucleotide 16 to nucleotide 32 on SEQ ID No. 5 for SEQ ID No. 26
- from nucleotide 503 to nucleotide 485 on SEQ ID No. 5 for SEQ ID No. 27
- from nucleotide 160 to nucleotide 176 on SEQ ID No. 11 for SEQ ID No. 28
- from nucleotide 1993 to nucleotide 1976 on SEQ ID No. 5 for SEQ ID No. 29
- from nucleotide 263 to nucleotide 280 on SEQ ID No. 11 for SEQ ID No. 30
- from nucleotide 1943 to nucleotide 1926 on SEQ ID No. 5 for SEQ ID No. 31
- from nucleotide 128 to nucleotide 145 on the nucleotide sequence depicted in Figure 22 for SEQ ID No. 32
- from nucleotide 1167 to nucleotide 1149 on SEQ ID No. 5 for SEQ ID No. 33
- from nucleotide 928 to nucleotide 911 on SEQ ID No. 5 for SEQ ID No. 34
- from nucleotide 677 to nucleotide 659 on SEQ ID No. 5 for SEQ ID No. 35
- from nucleotide 1605 to nucleotide 1587 on SEQ ID No. 5 for SEQ ID No. 36
- from nucleotide 1 to nucleotide 18 on the nucleotide sequence depicted in Figure 13 for SEQ ID No. 37
- from nucleotide 833 to nucleotide 813 on th nucleotide sequence depicted in Figure 13 for SEQ ID No. 38
- from nucleotide 25 to nucleotide 42 on the nucleotide sequence depicted in Figure 13 for SEQ ID No. 39
- from nucleotide 506 to nucleotide 488 on the nucleotide sequence depicted in Figure 13 for SEQ ID No. 40

The nucleotide sequences according to the invention can have, moreover, uses in gene th rapy, in particular for controlling the phenomena of apoptosis and

of reversion of transformation.

The nucleotide sequences according to the invention may, moreover, be used for th production of recombinant SR-p70 proteins, according to the definition which has been given to this term.

These proteins may be produced from the nucleotide sequences defined above, according to techniques of production of recombinant products which are known to a person skilled in the art. In this case, the nucleotide sequence used is placed under the control of signals permitting its expression in a cell host.

An effective system for production of a recombinant protein necessitates having at one's disposal a vector, for example of plasmid or viral origin, and a compatible host cell.

The cell host may be selected from prokaryotic systems such as bacteria, or eukaryotic systems such as, for example, yeasts, insect cells, CHO cells (Chinese hamster ovary cells) or any other system advantageously available. A preferred cell host for the expression of proteins of the invention consists of the *E. coli* bacterium, in particular the strain MC 1061 (Clontec).

The vector must contain a promoter, translation initiation and termination signals and also the appropriate transcription regulation regions. It must be capable of being maintained stably in the cell and can, where appropriate, possess particular signals specifying the secretion of the translated protein.

These various control signals are selected in accordance with the cell host used. To this end, the nucleotide sequences according to the invention may b inserted into vectors which are autonomously replicating within the selected host, or vectors which are integrative for the chosen host. Such vectors will be pr par d according to m thods commonly used by a person skilled in the art, and th clones resulting therefrom may be introduced into a suitable host by standard m thods such as, for exampl , l ctroporation.

The cloning and/or expr ssion vectors containing

at least one of the nucleotide sequences defined above also form part of the present invention.

A preferred cloning and expression vector is the plasmid pSE1, which contains the elements necessary for its use both as a cloning vector in *E. coli* (origin of replication in *E. coli* and ampicillin resistance gene originating from the plasmid pTZ 18R) and as an expression vector in animal cells (promoter, intron, polyadenylation site, origin of replication of the SV40 virus), as well as the elements enabling it to be copied as a single strand with the object of sequencing (origin of replication of phage f1).

The characteristics of this plasmid are described in Application EP 0,506,574.

Its construction and also the integration of the cDNAs originating from the nucleic acid sequences of th invention are, moreover, described in the examples below.

According to a preferred embodiment, the proteins of the invention are in the form of fusion proteins, in particular in the form of a protein fused with glutathione S-transferase (GST). A designated expression vector in this case is represented by the plasmid vector pGEX-4T-3 (Pharmacia ref-27.4583).

The invention relates, in addition, to the host cells transfected by these aforementioned vectors. Thes cells may be obtained by introducing into host cells a nucleotide sequence inserted into a vector as defined above, followed by culturing of the said cells under conditions permitting the replication and/or expression of the transfected nucleotide sequence.

These cells are usable in a method of production of a recombinant polypeptide of sequence SEQ ID No. 2, SEQ ID No. 4, SEQ ID No. 6, SEQ ID No. 8, SEQ ID No. 10, SEQ ID No. 12, SEQ ID No. 14, SEQ ID No. 16 or SEQ ID No. 18 or any biologically active fragment or derivative of th latter.

The method of production of a polyp ptide of the invention in recombinant form is its lf included in the present invention, and is charact rized in that the

transfected cells are cultur d under conditions permitting the expression of a recombinant polypeptide of sequence SEQ ID No. 2, SEQ ID No. 4, SEQ ID No. 6, SEQ ID No. 8, SEQ ID No. 10, SEQ ID No. 12, SEQ ID No. 14, SEQ ID No. 16 or SEQ ID No. 18 or of any biologically active fragment or derivative of the latter, and in that the said recombinant polypeptide is recovered.

The purification methods used are known to a person skilled in the art. The recombinant polypeptide may be purified from lysates and cell extracts or from the culture medium supernatant, by methods used individually or in combination, such as fractionation, chromatographic methods, immunoaffinity techniques using specific mono- or polyclonal antibodies, and the like. A preferred variant consists in producing a recombinant polypeptide fused to a "carrier" protein (chimeric protein). The advantage of this system is that it permits a stabilization and a decrease in proteolysis of the recombinant product, an increase in solubility during *in vitro* renaturation and/or a simplification of the purification when the fusion partner possesses an affinity for a specific ligand.

Advantageously, the polypeptides of the invention are fused with glutathione S-transferase at the N-terminal position (Pharmacia "GST" system). The fusion product is, in this case, detected and quantified by means of the enzyme activity of the GST. The colorimetric reagent used is a glutathione acceptor, a substrate for GST. The recombinant product is purified on a chromatographic support to which glutathione molecules have been coupled beforehand.

The mono- or polyclonal antibodies capable of specifically recognizing an SR-p70 protein according to the definition given above also form part of the invention. Polyclonal antibodies may b obtained from the serum of an animal immunized against protein, produced, for example, by g netic recombination according to the method describ d above, according to standard pr cedures.

The monoclonal antibodies may be obtained

according to the traditional hybridoma culture method describ d by Köhler and Milstein, Nature, 1975, *256*, 495-497.

Advantageous antibodies are antibodies directed against the central region lying between residue 110 and residue 310 for the sequences SEQ ID No. 2 or 6, or between residue 60 and residue 260 for the sequence SEQ ID No. 8.

The antibodies according to the invention are, for example, chimeric antibodies, humanized antibodies or Fab and $F(ab')_2$ fragments. They may also take the form of immunoconjugates or labelled antibodies.

Moreover, besides their use for the purification of the recombinant polypeptides, the antibodies of the invention, especially the monoclonal antibodies, may also be used for detecting these polypeptides in a biological sample.

Thus they constitute a means of immunocytochemical or immunohistochemical analysis of the expression of SR-p70 proteins on sections of specific tissues, for example by immunofluorescence, gold labelling or enzyme immunoconjugates.

They make it possible, in particular, to demonstrate an abnormal accumulation of SR-p70 proteins in certain tissues or biological samples, which makes them useful for detecting cancers or monitoring the progression or remission of pre-existing cancers.

More generally, the antibodies of the invention may be advantageously employed in any situation where the expression of an SR-p70 protein has to be observed.

Hence the invention also relates to a method of *in vitro* diagnosis of pathologies correlated with an expression or an abnormal accumulation of SR-p70 proteins, in particular the phenomena of carcinogenesis, from a biological sample, characterized in that at least one antibody of th invention is brought into contact with th said bi logical sample under conditions p rmitting the possible formation of specific immunological complex s betwe n an SR-p70 protein and the said

antibody or antibodies, and in that the specific immunological complexes possibly formed are detected.

The invention also relates to a kit for the *in vitro* diagnosis of an abnormal expression or accumulation of SR-p70 proteins in a biological sample and/or for measuring the level of expression of this protein in the said sample, comprising:

- at least one antibody specific for an SR-p70 protein, optionally bound to a support,
- means of visualization of the formation of specific antigen-antibody complexes between an SR-p70 protein and the said antibody, and/or means of quantification of these complexes.

The invention also relates to a method of early diagnosis of tumour formation, by detecting autoantibodies directed against an SR-p70 protein in an individual's serum.

Such a method of early diagnosis is characterized in that a serum sample drawn from an individual is brought into contact with a polypeptide of the invention, optionally bound to a support, under conditions permitting the formation of specific immunological complexes between the said polypeptide and the autoantibodies possibly present in the serum sample, and in that the specific immunological complexes possibly formed are detected.

A subject of the invention is also a method of determination of an allelic variability, a mutation, a deletion, an insertion, a loss of heterozygosity or a genetic abnormality of the SR-p70 gene which may be involved in pathologies, characterized in that it utilizes at least one nucleotide sequence described above. Among the methods of determination of an allelic variability, a mutation, a deletion, an insertion, a loss of heterozygosity or a genetic abnormality of the SR-p70 gene, preference is given to the method which is characterized in that it comprises at least one step of PCR amplification of the target nucleic acid sequence of SR-p70 liable to exhibit a polymorphism, a mutation, a

deletion or an insertion, using a pair of primers of nucl otide sequences defined above, a step during which the amplified products are treated using a suitable restriction enzyme and a step during which at least one of the products of the enzyme reaction is detected or assayed.

The invention also comprises pharmaceutical compositions comprising as active principle a polypeptide corresponding to the above definitions, preferably in soluble form, in combination with a pharmaceutically acceptable vehicle.

Such compositions afford a novel approach to treating the phenomena of carcinogenesis at the level of the control of multiplication and cell differentiation.

Preferably, these compositions can be administered systemically, preferably intravenously, intramuscularly, intradermally or orally.

Their optimal modes of administration, dosages and pharmaceutical dosage forms may be determined according to the criteria generally borne in mind in establishing a therapeutic treatment suitable for a patient, such as, for example, the patient's age or body weight, the severity of his or her general state, the tolerability of treatment and the observed side effects, and the like.

Lastly, the invention comprises a method of gene therapy, in which nucleotide sequences coding for an SR-p70 protein are transferred to target cells by means of inactivated viral vectors.

Other features and advantages of the invention are to be found in the remainder of the description, with the examples and the figures for which the legends are given below.

LEGEND TO THE FIGURES

Figur 1: Nucleic acid comparison of monkey SR-p70a cDNA (corresponding to nucleotides 1-1549 of SEQ ID No. 1) with the nucleic acid sequence of monkey p53 cDNA (SEQ ID NO. 43)

Figur 2: Protein comparison of monkey SR-p70a (amino acids 1-450 of SEQ ID No. 1) with monkey p53 protein (SEQ ID NO. 44) (sw: p53-cerae).

Figure 3: Comparison of the nucleic acid sequence of monkey SR-p70a and b cDNA (corresponding, respectively, to SEQ ID No. 1 and SEQ ID No. 3).

Figure 4: Nucleic acid sequence (SEQ ID NO. 1) and deduced protein sequence (SEQ ID NO. 2) of monkey SR-p70a.

Figure 5: Partial nucleic acid sequence (SEQ ID NO. 3) and complete deduced protein sequence (SEQ ID NO. 4) of monkey SR-p70b.

Figure 6: Partial nucleic acid sequence (SEQ ID NO. 5) and deduced complete protein sequence (SEQ ID NO. 6) of human SR-p70a ~~(corresponding to SEQ ID No. 5)~~.

Figure 7: Partial nucleic acid sequence (SEQ ID NO. 7) and complete deduced protein sequence (SEQ ID NO. 8) of mouse SR-p70c ~~(corresponding to SEQ ID No. 7)~~.

Figure 8: Partial nucleic acid sequence (SEQ ID NO. 9) and partially deduced protein sequence (SEQ ID NO. 10) of mouse SR-p70a ~~(corresponding to SEQ ID No. 9)~~.

Figure 9: Multialignment of the proteins deduced from monkey (a and b), human (a) (SEQ ID NO. 6), and mouse (a and c), SR-p70 cDNAs.

Figure 10a: Immunoblot of the SR-p70 protein.

Figure 10b: Detection of the endogenous SR-p70 protein.

Figur 11: Chromosomal localization of th human SR-p70 gene. The signal appears on chromosome 1, in the p36 r gion.

Figure 12: Genomic structure of the SR-p70 g ne and

comparison with that of the p53 gene. The human protein sequences of SR-p70a (SEQ ID NO.6) (upper line of the alignment) and of p53 (SEQ ID NO.45) (lower line) are divided up into peptides on the basis of the respective exons from which they are encoded. The figures beside the arrows correspond to the numbering of the corresponding exons.

Figure 13: Human genomic sequence of SR-p70 from the 3' end of intron 1 to the 5' end of exon 3 (SEQ ID NO.46). The introns are boxed. At positions 123 and 133, two variable nucleic acid positions are localized (G → A at 123 and C → T at 133). The restriction sites for the enzyme StyI are underlined (position 130 in the case where a T is present instead of a C at position 133, position 542 and position 610). The arrows indicate the positions of the nucleic acid primers used in Example XI.

Figure 14: Nucleic acid comparison of the 5' region of the human cDNAs of SR-p70d (SEQ ID NO.12) and of SR-p70a (SEQ ID NO.5).

Figure 15: Multialignment of the nucleic acid sequences corresponding to human SR-p70a, b, d, e, and f.

Figure 16: Multialignment of the proteins deduced from human SR-p70 (a, b, d, e and f) cDNAs.

Figure 17: Partial nucleic acid sequence (SEQ ID NO.94) (SEQ ID NO.5) and partial deduced protein sequence (SEQ ID NO.47) of human SR-p70a. The two bases in bold characters correspond to two variable positions (see Figure 6). This sequence possesses a more complete non-coding 5' region than the one presented in Figure 6.

Figure 18: Analysis of the SR-p70a transcripts after PCR amplification.
lane M: 1 kb ladder (GIBCO-BRL) molecular weight markers
lane 1: line HT29
lane 3: line SK-N-AS
lane 5: line UMR-32
lane 7: line U-373 MG
lane 9: line SW 480
lane 11: line CHP 212
lane 13: line SK-N-MC

lanes 2, 4, 6, 8, 10, 12, 14: negative controls corresponding to lanes 1, 3, 5, 7, 9, 11 and 13, respectively (absence of inverse transcriptase in the RT-PCR reaction).

Figure 19: A: Analysis by agarose gel electrophoresis of genomic fragments amplified by PCR (from the 3' end of intron 1 to the 5' end of exon 3). The numbering of the lanes corresponds to the numbering of the control population. Lane M: molecular weight markers (1 kb ladder).

B: Analysis identical to that of part A, after digestion of the same samples with the restriction enzym StyI.

Figure 20: Diagrammatic representation with a partial restriction map of the plasmid pCDNA3 containing human SR-p70a.

EXAMPLE I

Cloning of SR-p70 cDNA from COS-3 cells

1. Culturing of COS-3 cells

COS-3 cells (African green monkey kidney cells transformed with the SV 40 virus T antigen) are cultured in DMEM medium (GIBCO-BRL reference 41 965-047) containing 2 mM L-glutamine and supplemented with 50 mg/l of gentamicin and 5% of foetal bovine serum (GIBCO-BRL reference 10231-074) to semi-confluence.

2. Preparation of the messenger RNA

a) Extraction of the messenger RNA

The cells are recovered in the following manner:

- the adherent cells are washed twice with PBS buffer (phosphate buffered saline, reference 04104040-GIBCO-BRL), then scraped off with a rubber scraper and centrifuged.

The cell pellet is suspended in the lysis buffer of the following composition: 4 M guanidine thiocyanate; 25 mM sodium citrate pH 7; 0.5% sarcosyl; 0.1 M β-mercaptoethanol. The suspension is sonicated using an Ultra-Turrax No. 231256 sonicator (Janke and Kundel) at maximum power for one minute. Sodium acetate pH 4 is added to a concentration of 0.2 M. The solution is extracted with one volume or a phenol/chloroform (5/1 v/v) mixture. The RNA contained in the aqueous phase is precipitated at -20°C using one volume of isopropanol. The pellet is resuspended in the lysis buffer. The solution is extracted again with a phenol/chloroform mixture and the RNA is precipitated with isopropanol. After washing of the pellet with 70% and then 100% ethanol, the RNA is resuspended in water.

b) Purification of the poly(A)$^+$ fraction of the RNA

Purification of poly(A)$^+$ fraction of th RNA is carried out using the DYNAL Dynabeads oligo(dT)$_{25}$ kit (reference 610.05) according to the protocol

recommended by the manufacturer. The principle is based on the use of superparamagnetic polystyren beads to which an oligonucleotid poly(dT)$_{25}$ is attached. The poly(A)$^+$ fraction of the RNA is hybridized with the oligo(dT)$_{25}$ coupled to the beads, which are trapped on a magnetic support.

3. *Production of the complementary DNA library*

a) Preparation of the complementary DNA

From 0.5 µg of the poly(A)$^+$ RNA from COS-3 cells obtained at the end of step 2, the [^{32}P]dCTP-labelled single-stranded complementary DNA is prepared (the complementary DNA obtained possesses a specific activity of 3000 dpm/ng) with the synthetic primer of the following sequence (comprising a BamHI site):

5'<GATCCGGGCC CTTTTTTTTT TTT<3' (SEQ ID NO.4)

in a volume of 30 µl of buffer of composition: 50 mM Tris-HCl pH 8.3, 6 mM $MgCl_2$, 10 mM DDT, 40 mM KCl, containing 0.5 mM each of the deoxynucleotid triphosphates, 30 µCi of [α-^{32}P]dCTP and 30 U of RNasin (Promega). After one hour of incubation at 37°C, then 10 minutes at 50°C, then 10 minutes again at 37°C, with 200 units of the enzyme reverse transcriptase RNase H$^-$ (GIBCO-BRL reference 8064A), 4 µl of EDTA are added.

b) Alkaline hydrolysis of the RNA template

6 µl of 2N NaOH solution are added and the mixture is then incubated for 5 minutes at 65°C.

c) Purification on a Sephacryl S-400 column

In order to remove the synthetic primer, the complementary DNA is purified on a column of 1 ml of Sephacryl S-400 (Pharmacia) equilibrated in TE buffer.

The first two radioactive fractions ar pooled and precipitated with 1/10 volume of 10 M ammonium acetate solution and 2.5 volumes of ethanol, this being done after extraction with one volume of chloroform.

d) Homopolymer addition of dG

The complementary DNA is elongated at the 3' end with a dG tail with 20 units of the enzym terminal transferase (Pharmacia 27073001). The mixture is incubated in 20 µl of buffer of composition: 30 mM Tris-HCl pH 7.6, 1 mM cobalt chloride, 140 mM cacodylic acid, 0.1 mM DTT, 1 mM dGTP, for 15 minutes at 37°C, and 2 µl of 0.5 M EDTA are then added.

e) Steps b) and c) are repeated again

f) Pairing of the cloning vector pSE1 (EP 506,574) and the complementary DNA in the presence of the adaptor.

The mixture is centrifuged, the pellet is dissolved in 33 µl of TE buffer, 5 µl (125 ng) of cloning vector pSE1, 1 µl (120 ng) of the adaptor of the following sequence (comprising an ApaI site):

5'AAAAAAAAAAAAAGGGCCCG3' (SEQ ID NO.48)

and 10 µl of 200 mM NaCl solution are added, and the reaction mixture is incubated for 5 minutes at 65°C and then allowed to cool to room temperature.

g) Ligation

The cloning vector and the single-stranded cDNA ar ligated in a volume of 100 µl with 32.5 units of the enzyme phage T4 DNA ligase (Pharmacia reference 270 87002) overnight at 15°C in a buffer of composition: 50 mM Tris-HCl pH 7.5, 10 mM $MgCl_2$, 1 mM ATP.

h) Synthesis of the second strand of the cDNA

The proteins are removed by phenol extraction followed by chloroform extraction, and 1/10 volume of 10 mM ammonium acetate solution and then 2.5 volumes of ethanol are then added. The mixture is centrifuged, the pellet is dissolved in a buffer of composition 33 mM Tris-acetate pH 7.9, 62.5 mM potassium acetate, 1 mM magn sium ac tat and 1 mM dithiothr itol (DTT), and the second strand of complementary DNA is synthesized in a volum of 30 µl with 30 units of the nzyme phage T4 DNA polymerase (Pharmacia reference 270718) and a

mixture of 1 mM the four deoxynucleotide triphosphates dATP, dCTP, dGTP and dTTP as well as two units of phage T4 gene 32 protein (Pharmacia reference 27-0213) for one hour at 37°C. The mixture is extracted with phenol and the traces of phenol are removed with a column of polyacrylamide P10 (Biogel P10-200-400 mesh - reference 15011050 - Biorad).

i) Transformation by electroporation

E. coli MC 1061 cells are transformed with the recombinant DNA obtained above by electroporation using a Biorad Gene Pulser apparatus (Biorad) used at 2.5 kV under the conditions specified by the manufacturer, and the bacteria are then grown for one hour in the medium known as LB medium (Sambrook *op. cit.*) of composition: bactotryptone 10 g/l; yeast extract 5 g/l; NaCl 10 g/l.

The number of independent clones is determined by plating out a 1/1000 dilution of the transformation after the first hour of incubation on a dish of LB medium with the addition of 1.5% of agar (w/v) and 100 µg/ml of ampicillin, hereinafter referred to as LB agar medium. The number of independent clones is 1 million.

j) Analysis of the cDNAs of the library

In the context of the analysis of individual clones of the library by nucleic acid sequencing of the 5' region of the cDNAs, one clone, designated SR-p70a, was shown to exhibit a partial homology with the cDNA of the already known protein, the p53 protein (Genbank X 02469 and X 16384) (Figure 1). The sequences were produced with the United States Biochemical kit (reference 70770) and/or the Applied Biosystems kit (references 401434 and/or 401628), which use the method of Sanger et al., Proc. Natl. Acad. Sci. USA; 1977, *14*, 5463-5467. The plasmid DNA is prepared from the WIZARD minipreparation kit (Promega reference A7510). The primers used are 16- to 22-mer oligonucleotides, complementary either to

the vector pSE1 in the region immediately at the 5' end of the cDNA, or to the sequence of the cDNA.

A second cDNA was isolated from the same library by screening, in a manner similar to the technique described in EXAMPLE III.3) below, with a fragment of SR-p70a the DNA labelled with ^{32}P with the BRL "Random Primers DNA labelling systems" kit (reference 18187-013). The hybridization and washing buffers are treated by adding 50% of formamide. The last wash is carried out in 0.1 × SSC/0.1% SDS at 60°C. This second sequence (SR-p70b cDNA) is identical to the first but an internal fragment has been deleted from it (Figure 3).

The two SR-p70 cDNAs, of length 2874 nucleotides (SR-p70a) and 2780 nucleotides (SR-p70b), correspond to the products of a single gene, an alternative splicing bringing about a deletion of 94 bases between nucleotides 1637 and 1732 and a premature termination of the corresponding encoded protein. The proteins deduced from the two cDNAs possess 637 amino acids and 499 amino acids, respectively (Figures 4 and 5).

EXAMPLE II

Obtaining of the sequence and cloning of the cDNA of the SR-p70a protein from HT-29 (human colon adenocarcinoma) cells

1) Culturing of HT-29 cells

The cells are cultured in McCoy's 5 medium (GIBCO 26600-023) with the addition of 10% of foetal calf serum (GIBCO 10081-23) and 50 mg/l of gentamicin, to semi-confluence.

2) Preparation of th complem ntary DNA

Th mess nger RNA is prepared as d scribed in EXAMPLE I.2. The cDNA is pr pared in a manner similar to that describ d in EXAMPLE I.3, with 5 μg of total messeng r RNA, using a poly$(T)_{12}$ primer. The reaction is

not interrupted with EDTA.

3) Specific amplification of the human cDNA by the so-called PCR technique

The polymerization is carried out with 4 µl of cDNA in 50 µl final with the buffer of the following composition: 10 mM Tris-HCl pH 8.3, 2.5 mM $MgCl_2$, 50 mM KCl in the presence of 10% DMSO, 0.5 mM dNTP, 4 µg/ml of each of the two nucleic acid primers and 2.5 units of TAQ DNA polymerase (Boehringer). The primer pairs were selected on the basis of the nucleic acid sequence of the COS-3 SR-p70 clone, in particular upstream of the translation initiation ATG and downstream of th translation stop TGA, and are of the following compositions:

sense primer: ACT <u>GGT ACC</u> GCG AGC TGC CCT CGG AG (SEQ ID No.49)
Kpn I restriction site

antisense primer: GAC <u>TCT AGA</u> GGT TCT GCA GGT GAC TCA G (SEQ ID No.50)
Xba I restriction site

The reaction is carried out for 30 cycles of 94°C/1 minute, 54-60°C/1 minute 30 seconds and 72°C/1 minute 30 seconds, followed by a final cycle of 72°C/6 minutes.

4) Obtaining of the sequence of the human cDNA

In a first step, the PCR product is removed from the oligonucleotides on a column of Sephacryl S-400, and then desalted by exclusion chromatography on a column of polyacrylamide P10 (Biorad reference 1504144). The sequencing reactions are carried out using the Applied Biosystems kit (reference 401628) with oligonucleotides specific for th cDNA. Th sequence obtained is very similar to that of monkey SR-p70a, and th deduced prot in contains 636 amino acids (Figure 6).

In a similar manner, other s quences originating from human lines or tissu s were obtained for the coding

portion of human SR-p70, in particular from the lung or pancreas. The prot ins deduced from these sequences are identical to thos obtained for the HT-29 line.

5) Cloning of the human cDNA into plasmid pCDNA3 (Invitrogen V 790-20)

The PCR product obtained in 3) and also the plasmid are digested with the two restriction enzymes Kpn I and Xba I and then purified after migration on a 1% agarose gel using the Geneclean kit (Bio 101 reference 3105). After ligation with 100 ng of insert and 10 ng of vector and transformation (technique described in EXAMPLE I.3.g and i), the recombinant clones are verified by sequencing using the Applied Biosystems kit mentioned above.

EXAMPLE III

Cloning of mouse SR-p70 cDNA from AtT-20 (pituitary tumour) cells

1) Cell culturing of the line AtT-20

The cells are cultured in Ham F10 medium (GIBCO 31550-023) with the addition of 15% of horse serum (GIBCO 26050-047), 2.5% of foetal calf serum (GIBCO 10081-073) and 50 mg/l of gentamicin, to semi-confluence.

2) Preparation of the complementary DNA library

The library is produced as described in EXAMPLE I. 2 and 3 from the cells cultured above.

3) Screening of the library

a) Preparation of the membranes

The clones of the library are plated out on LB agar medium (Petri dishes 150 mm in diameter) coated with Biodyne A membranes (PALL ref rence BNNG 132). After on night at 37°C, the clones are transferr d by contact onto fr sh membranes. The latter are treated by d positing them on 3 mm Whatman paper soaked with the following solutions: 0.5 N NaOH, 1.5 M NaCl for 5 minutes, then

0.5 M Tris-HCl pH 8, 1.5 M NaCl for 5 minutes. After treatment with proteinase K in the following buffer: 10 mM Tris-HCl pH 8, 10 mM EDTA, 50 mM NaCl, 0.1% SDS, 100 µg/ml proteinase K, for one hour at room temperature, the membranes are washed copiously in 2 × SSC (sodium citrate, NaCl), dried and then incubated in an oven under vacuum at 80°C for 20 minutes.

b) Preparation of the probe

On the basis of monkey and human SR-p70 cDNA sequences, a first sequence was produced on a fragment amplified from line AtT-20 mRNA as described in EXAMPLE II.3 and 4, with the oligomers of the following compositions:

sense primer: GCC ATG CCT GTC TAC AAG (SEQ ID NO. 22)

antisense primer: ACC AGC TGG TTG ACG GAG (SEQ ID NO. 23)

On the basis of this sequence, an oligomeric probe specific for mouse was chosen and possesses the following composition: (SEQ ID NO. 57)

GAG CAT GTG ACC GAC ATT G.

100 ng of the probe are labelled at the 3' end with 10 units of terminal transferase (Pharmacia) and 100 µCi of [α-^{32}P]dCTP 3000 Ci/mmol (Amersham reference PB 10205) in 10 µl of the following buffer: 30 mM Tris-HCl pH 7.6, 140 mM cacodylic acid, 1 mM $CoCl_2$, 0.1 mM DTT for 15 minutes at 37°C. The radiolabelled nucleotides not incorporated are removed on a column of polyacrylamide P10 (Biorad, reference 1504144). The probe obtained has a specific activity of approximately 5 × 10^8 dpm/µg.

c) Prehybridization and hybridization

The membranes prepared in a) are prehybridized for 30 minutes at 42°C in 6 × SSC, 5 × Denhart's, 0.1% SDS, and then hybridized for a few hours in the same buffer with the addition of the probe prepared in b) in the proportion of 10^6 dpm/ml.

d) Washing and exposure of the membran s

The membranes are washed twice at room temperature in 2 × SSC/0.1% SDS buffer and then for one hour at 56°C in 6 × SSC/0.1% SDS. The hybridized clones are visualized with KODAK XOMAT films. A positive clone

containing the mouse SR-p70 is selected and hereinafter designated as SR-p70c.

4) Sequencing of mouse SR-p70 and analysis of the sequence

The sequence is obtained using the Applied Biosystem kit (reference 401628). The protein sequence deduced from mouse SR-p70c cDNA (Figure 7) exhibits a very strong homology with the human and monkey sequences, except in the N-terminal portion which diverges strongly (see Figure 9). Using the so-called PCR technique in a similar manner to that described in EXAMPLE II.3 and 4, a second 5′ sequence (originating from the same AtT-20 library) was obtained (Figure 8). The deduced N-terminal protein sequence (sequence designated SR-p70a) is very similar to that deduced from human and monkey SR-p70 cDNAs (SR-p70a) (Figure 9). The line AtT-20 hence affords at least two SR-p70 transcripts. The latter 2 diverge in the N-terminal portion through different splicings.

EXAMPLE IV

1) Production of recombinant SR-p70 protein in E. coli

a) Construction of the expression plasmid

This consists in placing the COOH-terminal portion of the monkey SR-p70a protein, from the valine at position 427 to the COOH-terminal histidine at position 637, in fusion with the glutathione S-transferase (GST) of the plasmid vector pGEX-4T-3 (Pharmacia reference 27-4583). For this purpose, the corresponding insert of SR-p70a (position 1434 to 2066) was amplified by PCR with 10 ng of plasmid containing monkey SR-p70a cDNA. The nucleic acid primers are of the following composition:

A (SEQ ID No. 52)

sense primer: TTT <u>GGA TCC</u> GTC AAC CAG CTG GTG GGC CAG

BamHI restriction sit

A (SEQ ID No. 53)

antisense primer: AAA <u>GTC GAC</u> GTG GAT CTC GGC CTC C

Sal I site

The fragment obtained and also th vector are dig sted with th restriction enzymes BamHI and Sal I and cloning is carried out as described in EXAMPLE II.5. The selected clone is referred to as pG SR-p70.

b) Expression and purification of the GST-pSR-p70 fusion protein

This step was carried out using the "bulk GST purification module" kit (Pharmacia Reference 27-4570-01).

In outline, the recombinant clone was cultured at 37°C in one litre of 2 × YTA medium + 100 µg/ml ampicillin. At OD 0.8, expression is induced with 0.5 mM IPTG for 2 hours at 37°C. After centrifugation, the cell pellet is taken up in cold PBS and then sonicated by ultrasound. After the addition of 1% Triton X-100, the preparation is incubated for 30 minutes with agitation at room temperature. After centrifugation at 12,000 g for 10 minutes at 4°C, the supernatant is recovered. Purification is then carried out on a glutathione-Sepharose 4B affinity chromatography column. Binding and washing are carried out in PBS buffer and elution is carried out by competition with reduced glutathione. The final concentration is brought to 300 µg/ml of fusion protein.

2) Production of SR-p70a protein in COS-3 cells

COS-3 cells are transfected with pSE1 plasmid DNA into which monkey SR-p70a cDNA has been cloned (EXAMPLE I.1), or with the vector pSE1 plasmid DNA as control, by the DEAE-dextran technique: the COS-3 cells are inoculated at 5×10^5 cells per 6 cm dish in culture medium containing 5% of foetal bovine serum (EXAMPLE I.1). After culture, the cells are rinsed with PBS. 1 ml of the following mixture is added: medium containing 6.5 µg of DNA, 250 µg/ml of DEAE-dextran and 100 µM chloroquine. The cells ar incubated at 37°C in 5% CO_2 for 4 to 5 hours. Th medium is aspirated off, 2 ml of PBS containing 10% of DMSO are added and the cells are incubated for one minut , shaking the dishes gently. The

medium is aspirated off again and the cells are rinsed twic with PBS. The cells are then incubated at 37°C with medium containing 2% of foetal bovine serum for the period during which expression takes place, which is generally 3 days.

The SR-p70a protein is then analysed as described in EXAMPLE IV by immunoblotting.

EXAMPLE V

Preparation of specific antibodies

150 µg of proteins of the sample prepared according to EXAMPLE IV were used to immunize a rabbit (New Zealand male weighing 1.5 to 2 kg approximately). The immunizations were performed every 15 days according to the protocol described by Vaitukaitis, Methods in Enzymology, 1981, 73, 46. At the first injection, one volume of antigenic solution is emulsified with one volume of Freund's complete adjuvant (Sigma reference 4258). Five boosters were administered in Freund's incomplete adjuvant (Sigma reference 5506).

EXAMPLE VI

Detection of the SR-p70 protein: Western immunoblotting

1) Materials used for immunoblotting

a) Cell lines used for immunoblotting

The following cell lines were cultured as described in the catalogue "Catalogue of cell lines and hybridomas, 7th edition, 1992" of the ATCC (American Typ Culture Collection): (COS-3, CV-1 (monkey kidney cell line), HT-29, U-373MG (human glioblastoma), MCF7 (human mammary adenocarcinoma), SKNAS (human neuroblastoma cultur d under the same conditions as COS-3), SK-N-MC (human neuroblastoma), IMR-32 (human neuroblastoma), CHP212 (human neuroblastoma cultured under the same conditions as CV-1), Saos-2 (osteosarcoma), SK-OV-3 (ovarian adenocarcinoma) and SW 480 (human colon adenocarcinoma).

b) COS-3 cells transfected by SR-p70a cDNA

COS-3 cells wer transfected as described in EXAMPLE IV.2. As a control, the cells were transfected with pSE1 plasmid DNA not containing recombinant SR-p70a cDNA.

2) Preparation of protein samples from a eukaryotic cell culture or from transfected cells

After culture, the cells are washed with PBS and then taken up in RIPA buffer (PBS with 1% NP40, 0.5% sodium deoxycholate, 0.5% SDS) supplemented with 10 µg/ml RNAse A, 20 µg/ml DNAse 1, 2 µg/ml aprotinin, 0.5 µg/ml leupeptin, 0.7 µg/ml pepstatin and 170 µg/ml PMSF. Th cells are sonicated by ultrasound at 4°C and left for 30 minutes at 4°C. After microcentrifugation at 12,000 rpm, the supernatant is recovered. The protein concentration is measured by the Bradford method.

3) Western blotting

5 or 50 µg of proteins (50 µg for the cell lines and 5 µg for transfected cells) are placed in 0.2 volume of the following 6 × electrophoresis buffer: 0.35 mM Tris-HCl pH 6.8, 10.3% SDS, 36% glycerol, 0.6 mM DTT, 0.012% bromophenol blue. The samples are applied and run in a 10% SDS-PAGE gel (30:0.8 Bis) and then electrotransferred onto a nitrocellulose membrane.

4) Visualization with the antibody

The membrane is incubated for 30 minutes in TBST blocking buffer (10 mM Tris-HCl pH 8, 150 mM NaCl, 0.2% Tween 20) with the addition of 5% of milk (GIBCO - SKIM MILK) at room temperature. The membrane is brought into contact successively with the anti-SR-p70 (αSR-p70) antibody in the same buffer for 16 hours at 4°C, washed 3 times for 10 minutes with TBST and then incubat d for one hour at 37°C with a second, anti-rabbit immunoglobulin antibody coupled t peroxidas (SIGMA A055). After thre washes of 15 minutes, the visualization is performed using the ECL kit (Amersham

RPN2106) by chemiluminescence.

In parall 1, the same samples were subjected to visualization with an anti-p53 (αp53) antibody (Sigma BP5312) followed by a second, anti-mouse immunoglobulin antibody.

5) Figures and results

Figure 10: Immunoblot of the SR-p70 protein

Figure 10a: Detection of the recombinant SR-p70 protein
- columns 1 and 3: COS-3 transfected by the vector pSE1.
- columns 2 and 4: COS-3 transfected by plasmid pSE1 containing SR-p70a cDNA.
- columns 1 and 2: visualization with the anti-SR-p70 (αSR-p70) antibody.
- columns 3 and 4: visualization with the anti-p53 (αp53) antibody.

Figure 10b: Detection of the endogenous SR-p70 protein
- columns 1: COS-3; 2: CV-1; 3: HT-29; 4: U-373 MG; 5: MCF7; 6: SKNAS; 7: SK-N-MC; 8: IMR-32; 9: CHP212; 10: Saos-2; 11: SK-OV-3 and 12: SW480.
A: Visualization with the αSR-p70 antibody
B: Visualization with the αp53 antibody.

The αSR-p70 antibody specifically recognizes th recombinant proteins (Figure 10a) and endogenous proteins (Figure 10b) and does not cross with p53. The analysis of human or monkey cell lines shows the SR-p70 protein, like p53, is generally weakly detectable. In contrast, when an accumulation of p53 exists, SR-p70 becomes, for its part also, more readily detectable (Figure 10b). A study by RT-PCR of the distribution of SR-p70 transcripts shows that the gene is expressed in all the cell types tested.

EXAMPLE VII

Cloning of the SR-p70 g ne and chrom somal localization

1) Cloning of SR-p70 gene

The library us d is a cosmid library prepared

in th EXAMPLE III.3, with an SR-p70 DNA fragment labelled with ^{32}P with the BRL "Random Primers DNA Labelling Systems" kit (reference 18187-013). The hybridization and washing buffers are treated by adding 50% of formaldehyde. The last wash is carried out in 0.1 x SSC/0.1% SDS at 60°C. In a similar manner, the SR-p70 gene was isolated from a library prepared with C57 black mouse genomic DNA.

An analysis and a partial sequencing of the clones demonstrate the presence of 14 exons with a structure close to that of the p53 gene, in particular in the central portion where the size and positioning of the exons are highly conserved (Figure 12). This structure was partially defined in mouse and in man.

As an example, the human genomic sequences of the 3' region of intron 1, of exon 2, of intron 3 and of the 5' region of exon 3 are presented in Figure 13.

2) Chromosomal localization of the SR-p70 gene in man

This was carried out with human SR-70 gene DNA using the technique described by R. Slim et al., Hum. Genet., 1991, *88*, 21-26. Fifty mitoses were analysed, more than 80% of which had double spots localized at 1p36 on both chromosomes and more especially at 1p36.2-1p36.3 (Figure 11). The identification of chromosome 1 and its orientation are based on the heterochromatin of th secondary constriction. The pictures were produced on a Zeiss Axiophot microscope, taken with a LHESA cooled CCD camera and treated with Optilab.

EXAMPLE VIII

A) Demonstration of an mRNA coding for a deduced human SR-p70 protein possessing both a shorter N-terminal end and a divergence.

1) Culturing of IMR-32 (human neuroblastoma) cells

The cells were cultured as described in the catalogue "Catalogue of cell lines and hybridomas, 7th edition, 1992" of the ATCC (American Type Culture Collection).

2) Preparation of the cDNA

The RNA is prepared as described in Example I.2.a. The cDNA is prepared in a manner similar to that described in Example I.3, with 5 µg total RNA in a final volume of 20 µl using a poly(T)$_{12}$ primer and with cold nucleotides. The reaction is not interrupted with EDTA.

3) Specific amplification of SR-p70 cDNA by the so-called PCR technique

The polymerization is carried out with 2 µl of cDNA in 50 µl final with the buffer of the following composition: 50 mM Tris-HCl pH 9.2, 16 mM $(NH_4)_2SO_4$, 1.75 mM $MgCl_2$, in the presence of 10% DMSO, 0.4 mM NTP, 100 ng of each of the two nucleic acid primers and 3.5 units of the mixture of TAQ and PWO polymerases (Boehringer Mannheim, ref. 1681 842).

The primer pair is of the following composition:

(SEQ ID No.26)
sense primer: AGGCCGGCGTGGGGAAG (position 16 to 32, Figure 6)
(SEQ ID No.27)
antisense primer: CTTGGCGATCTGGCAGTAG (position 503 to 485, Figure 6).

The reaction is carried out for 30 cycles at 95°C/30 seconds, 58°C/1 minute and 68°C/2 minutes 30 seconds, followed by a final cycle of 68°C/10 minutes.

The PCR product is subjected to electrophoresis on a 1% agarose gel (TAE buffer). After ethidium bromide staining, two major bands are revealed: a band approximately 490 bp in size (expected size (see Figure 6)) and an additional band approximately 700 bp in size. The latter is extracted from the gel using the "Geneclean" kit (Bio 101, ref 1001 400). After a desalting on a column of polyacrylamide P10 (Biorad, ref

15011050), the fragment is subjected to a further PCR amplification for 10 cycles as described above.

4) Determination of the sequence of the amplified product

In a first step, the PCR product is removed from the oligonucleotides on a column of Sephacryl S-400 (Pharmacia 17-0609-01) and then desalted on a column of P10. The sequencing reaction is carried out using the Applied Biosystems kit (ref. 401 628) (373 DNA sequencer) with the antisense primer.

The sequence obtained is identical to the SR-p70 cDNA sequence (Example II.4) with an insertion of 198 bp between positions 217 and 218 (Figure 14). The deduced N-terminal protein sequence (sequence designated SR-p70d) is 49 amino acids shorter, with a divergence of the first 13 amino acids (sequence ID No. 13). There is hence coexistence of at least two different SR-p70 transcripts as already described for the mouse AtT-20 line.

B) Cloning of human SR-p70 and demonstration of an mRNA coding for a deduced human SR-p70 protein possessing the same N-terminal end as SR-p70d and a divergence in the C-terminal portion

1) Specific amplification of SR-p70 cDNA by the so-called PCR technique

The amplification was carried out as described in EXAMPLE VIII.A from purified RNA of IMR-32 cells with the primer pair of the following composition:

sense primer: GCG GCC ACG ACC GTG AC$_A$ (position 160 to 176, sequence ID No. 11) (SEQ ID No. 28)

antisense primer: GGC AGC TTG GGT CTC TGG$_A$ (position 1993 to 1976, Figure 6). (SEQ ID No. 29)

After removal of the excess primers on an S400 column and desalting on a P10 column, 1 µl of th sample is subjected again to a PCR with th primer pair of the following composition:

sense primer: TAT CTC GAG CTG TAC GTC GGT GAC CCC
XhoI (position

263 to 280, sequence ID No. 11)
antisense primer: ATA TCT AGA TCA GTG GAT CTC GGC CTC (SEQ ID No. 9)
XbaI (SEQ ID NO 35) (position
1943 to 1926, Figure 6).

2) Cloning of the amplified product into plasmid pCDNA3

The PCR product obtained in 1) is desalted on a P10 column, digested with the restriction enzymes XhoI and XbaI and then cloned into plasmid pCDNA3 as described in EXAMPLE II.5. Two recombinant clones are sequenced using the Applied Biosystems kit with the oligonucleotides specific for SR-p70 cDNA.

The first sequence obtained corresponds to the complete sequence of the mRNA coding for SR-p70 described in EXAMPLE VIII.a. The deduced protein contains 587 amino acids (sequence ID No. 13 and Figure 16).

The second sequence obtained is identical to the SR-p70d cDNA sequence described above, but with two deletions, of 149 bp and of 94 bp between positions 1049 and 1050 on the one hand, and between positions 1188 and 1189 on the other hand (sequence ID No. 14 and Figure 15). The protein sequence deduced from this second sequence reveals a protein having an N-terminal portion 49 amino acids shorter, with a divergence in the first 13 amino acids as well as a divergence of protein sequence between amino acids 350 and 397 (sequence ID No. 15 and Figure 16) (sequence designated SR-p70e). The deduced protein contains 506 amino acids.

C) Demonstration of an mRNA coding for a deduced human SR-p70 protein possessing a shorter N-terminal end

1) Culturing of SK-N-SH (human neuroblastoma) cells

The cells are cultivated as described in the "Catalogue of cell lines and hybridomas, 7th edition, 1992" of the ATCC (American Type Culture Collection).

2) Preparation of the cDNA and amplification of SR-p70 cDNA by the so-called PCR technique

These steps are carried out as described in EXAMPLE VIII.A with the primer pair of the following composition:

sense primer: AGG GGA CGC AGC GAA ACC (position 128 to 145, Figure 17) (SEQ ID No. 37)

antisense primer: GGC AGC TTG GGT CTC TGG (position 1993 to 1976, Figure 6). (SEQ ID No. 29)

The sequencing is carried out with the Applied Biosystem kit with primers specific for SR-p70 cDNA, and reveals two cDNAs:

- a first cDNA corresponding to the mRNA coding for SR-p70a
- a second cDNA having a deletion of 98 bp between positions 24 and 25 (sequence ID No. 16 and Figure 15).

This deletion comprises the translation initiation ATG of SR-p70a. The protein deduced (designated SR-p70f) from this second cDNA possesses a translation initiation ATG downstream corresponding to an internal ATG of SR-p70a. The deduced protein hence contains 588 amino acids (sequence ID No. 17 and Figure 16) and is truncated with respect to the 48 N-terminal amino acids of SR-p70a.

D) Demonstration of an mRNA coding for human SR-p70b

1) Culturing of K562 cells

The cells are cultured as described in the "Catalogue of cell lines and hybridomas, 7th edition, 1992" of ATCC (American Type Culture Collection).

2) Preparation of the cDNA, amplification of SR-p70 cDNA by the so-called PCR technique and sequencing

Th se steps are carried out as described in EXAMPLE VIII.C.

The sequencing reveals two cDNAs:

A first cDNA corresponding to the mRNA coding for SR-p70a, and a second cDNA having a deletion of 94 bp

between positions 1516 and 1517 (sequence ID No. 18 and Figure 15). Th deduced protein (designated SR-p70b) contains 199 amino acids and possesses a C-terminal sequence truncated by 137 amino acids relative to SR-p70a, with the last 4 amino acids divergent (sequence ID No. 19 and Figure 21).

This cDNA is similar to the one described in EXAMPLE I relating to monkey SR-p70b.

The molecules described in this example (EXAMPLE VIII.A, B, C and D) reveal SR-p70 variants which are the outcome of differential splicings of the primary mRNA, transcribed by the SR-p70 gene.

The SR-p70a is encoded by an mRNA composed of 14 exons (see EXAMPLE VII). This is the reference protein. SR-p70b is the outcome of an insertion between exons 3 and 4 and of the absence of exons 11 and 13. SR-p70f is the outcome of the absence of exon 2. This example describes the existance of SR-p70 variants non-exhaustively, with a strong probability of existence of other variants. Similarly, the existence of these variants described in this example, as well as SR-p70a, is not limited to the lines in which they have been demonstrated. In effect, studies performed by RT-PCR showed that these variants are to be found in the various lines studied.

Furthermore, the initiation methionine of SR-p70f corresponds to an internal methionine of SR-p70a, suggesting the possibility of initiation downstream on the mRNA coding for SR-p70a.

EXAMPLE IX

Obtaining a 5′ sequence of human SR-p70a mRNA

1) Amplification of th 5′ nd of SR-p70 cDNA by PCR

The cell culturing and the preparations of total RNA and of cDNA are carri d out as described in EXAMPLE VIII.1 and 2. The RNA template is hydrolysed by incubation for 5 minutes at 65°C after th addition of 4 µl of 500 mM EDTA and 4 µl of 2 N NaOH. The sample is

then desalted on a P10 column. The cDNA is elongated at the 3' end with a dG tail as described in EXAMPLE I.3.d, in a final volume of 40 µl. After the addition of 4 µl of 500 mM EDTA and 4 µl of 2 N NaOH, the cDNA is incubated at 65°C for 3 minutes and then desalted on a P10 column. PCR amplification is carried out as described in EXAMPLE VIII.3 with 8 µl of cDNA and for 30 cycles with the primer pair of the following composition:

sense primer: C C C C C C C C C C C C C C N (where N equals G, A or T) (SEQ ID No.41)

antisense primer: CCATCAGCTCCAGGCTCTC (SEQ ID NO.33) (position 1167 to 1149, Figure 6).

After removal of the excess primers on an S-400 column and desalting on a P10 column, 1 µl of the sample is subjected again to a PCR with the pair of the following composition:

sense primer: C C C C C C C C C C C C C C N (SEQ ID NO.41)

antisense primer: CCAGGACAGGCGCAGATG (SEQ ID NO.34) (position 928 to 911, Figure 6).

The sample, passed again through an S-400 column and a P10 column, is subjected to a third amplification for 20 cycles with the following pair:

sense primer: C C C C C C C C C C C C C C N

antisense primer: CTTGGCGATCTGGCAGTAG (position 503 to 485, Figure 6).

2) Determination of the SR-p70 cDNA 5' sequence

The sequence is produced as described in EXAMPLE VIII.4. This sequence reveals a non-coding 5' region of at least 237 bases upstream of the initiation ATG of SR-p70a (Figure 17). By comparison of this sequence (obtained from the line IMR-32) with the one obtained from the line HT-29 in particular (Figure 6), two point differences (Figure 17: see bold characters) are revealed (G → A and C → T), position d, respectively, at -20 and -30 from the initiation ATG of SR-p70a (Figures 6 and 17). This variability is located in exon 2 (Figure 13). It is not ruled out that this variability is also to be found within a coding frame as the outcome of an

alternative splicing as described in EXAMPLES III in mouse and VIII in man, or alternatively as the outcome of a translation initiation on a CTG (as has been demonstrated for FGFb (Proc. Natl. Acad. Sci USA, 1989, 86, 1836 - 1840)).

Similarly, it is not ruled out that this variability has a repercussion on the translation of SR-p70 or on the splicing of the primary RNA.

At all events, this variability, probably of allelic origin, may serve as a marker, either at genomic level (see EXAMPLE XI) or at mRNA level (see EXAMPLE X).

EXAMPLE X

1) Analysis by PCR of the transcriptional expression of SR-p70a in cell samples (RT-PCR)

Cell culturing (SK-N-AS, SK-N-MC, HT-29, U-373MG, SW480, IMR-32, CHP212) is carried out as described in Example VI.1.a (referred to the catalogue "Catalogue of cell lines and hybridomas, 7th edition 1992" of the ATCC).

The preparation of the cDNA and the PCR amplification are carried out as described in EXAMPLE VIII.2 and 3. The primer pair used is of the following composition:

sense primer: AGGGGACGCAGCGAAACC (SEQ ID No. 32), (position 128 to 145, Figure 17)

antisense primer: GGCAGCTTGGGTCTCTGG (SEQ ID No. 29), (position 1993 to 1976, Figure 6).

The samples are analysed by electrophoresis on a 1% agarose gel and visualization with ethidium bromid (Figure 18).

The size of the band obtained in the samples corresponds to the xpected size (appr ximately 2 kb, Figures 6 and 17). The intensity of th bands obtained is reproducible. A reamplification of 1 µl of the sample under the same conditions for 20 cycles rev als a band in each of the sampl s.

2) Determination of the sequence of the amplified products

After passage of the samples through S-400 and P10 columns, sequencing is carried out on an Applied Biosystems sequencer 373 with the reference kit 401 628. The primers used are, inter alia, the following:

	position	Figure
AGGGGACGCAGCGAAACC (SEQ ID No. 32)	128 to 145	22
CTTGGCGATCTGGCAGTAG (SEQ ID No. 27)	503 to 485	6
GATGAGGTGGCTGGCTGGA (SEQ ID No. 35)	677 to 659	6
CCATCAGCTCCAGGCTCTC (SEQ ID No. 33)	1167 to 1149	6
TGGTCAGGTTCTGCAGGTG (SEQ ID No. 34)	1605 to 1587	6
GGCAGCTTGGGTCTCTGG (SEQ ID NO. 29)	1993 to 1976	6

No protein difference in the SR-p70a was detected. However, sequences obtained reveal a double variability at positions -20 and -30 upstream of the initiation ATG of SR-p70a (Figures 6 and 17). This variability, probably of allelic origin, enables two classes of transcripts to be defined: a first class possessing a G at position -30 and a C at position -20 (class G^{-30}/C^{-20}) and a second class possessing a difference at two positions: an A at -30 and a T at -20 (class A^{-30}/T^{-20}).

First class: SK-N-AS, SK-N-MC, HT-29, U-373MG, SW480.

Second class: IMR-32, CHP212.

EXAMPLE XI

Analytical method of determination of the allelic distribution of the SR-p70 gene in a population of 10 persons

This allelic distribution is based on the allelic variability demonstrated in EXAMPLES IX and X:

- G^{-30}/C^{-20} allele possessing, respectively, a G and a C at positions -30 and -20 upstream of the initiation ATG of SR-p70a.
- A^{-30}/T^{-20} allele possessing, respectively, an A and

a T at the same positions.

This variability may be demonstrated by the use of restriction enzymes that differentiate the two alleles (Figure 13). As an example:

- Enzyme Bpl I having a cleavage site only on the G^{-30}/C^{-20} allele in the zone of interest (this site encompasses both variable positions).
- Enzyme StyI having a cleavage site only on the A^{-30}/T^{-20} allele in the zone of interest.

1) Genomic amplification of exon 2 by PCR

The polymerization reaction is carried out with 500 ng of purified genomic DNA, in 50 µl final with the conditions described in Example VIII.3.

The primer pair is of the following position:

(SEQ ID NO. 37)

Sense primer: CACCTACTCCAGGGATGC (position 1 to 18, Figure 13)

(SEQ ID NO. 38)

Antisense primer: AGGAAAATAGAAGCGTCAGTC (position 833 to 813, Figure 13).

The reaction is carried out for 30 cycles as described in EXAMPLE VIII.3.

After removal of the excess primer on an S-400 column and desalting on a P10 column, 1 µl of the sample is amplified again for 25 cycles under the same conditions with the following primer pair:

(SEQ ID NO. 39)

Sense primer: CAGGCCCACTTGCCTGCC (position 25 to 32, Figure 13)

(SEQ ID NO. 40)

Antisense primer: CTGTCCCCAAGCTGATGAG (position 506 to 488, Figure 13).

The amplified products are subjected to electrophoresis on a 1% agarose gel (Figure 19-A).

2) Digestion with the restriction enzyme StyI

The samples are desalted beforehand on a P10 column and then digested with the restriction enzyme StyI (BRL 15442-015) in the buffer of the following composition: 50 mM Tris-HCl pH 8, 100 mM NaCl, 10 mM $MgCl_2$, at 37°C for 30 min. The digestion products are analysed by electrophor sis on a 1% agaros gel (TAE buffer). Visualization is carried ut by ethidium bromide

staining (Figure 19-B).

A band of 482 base pairs characterizes the G^{-30}/C^{-20} allele (Figures 13 and 19). The presence of a band of 376 base pairs and a band of 106 base pairs characterize the A^{-30}/T^{-20} allele (allele possessing a StyI cleavage site).

On the population of 10 persons, 2 persons exhibit the G^{-30}/C^{-20} and A^{-30}/T^{-20} alleles, the other 8 persons being homozygous with the G^{-30}/C^{-20} allele. The study of a fresh population of 9 persons demonstrated 3 heterozygous persons exhibiting the G^{-30}/C^{-20} and A^{-30}/T^{-20} alleles, the other 6 persons being homozygous for the G^{-30}/C^{-20} allele.

EXAMPLE XII

Test of reversion of transformation of the line SK-N-AS by transfection with SR-p70 cDNA

The expression vector used is described in EXAMPLE II.5 and shown diagrammatically in Figure 15. The method used is the so-called calcium phosphate method described by Graham et al. (Virology 1973, 54, 2, 536-539). The line is inoculated in the proportion of 5×10^5 cells per dish 6 cm in diameter in 5 ml of the medium described in Example I.1. The cells are cultured at 37°C and with 5% CO_2 overnight. The transfection medium is prepared in the following manner: the following mixture is prepared by adding, in order, 1 ml of HEBS buffer (8 mg/ml NaCl, 370 µg/ml KCl, 125 µg/ml $Na_2HPO_4.2H_2O$, 1 mg/ml dextrose, 5 mg/ml Hepes pH 7.05), 10 µg of the plasmid to be transfected and 50 µl of 2.5 M $CaCl_2$ added dropwise. The transfection medium is left for 30 min at room temperature and then added dropwise to the medium contained in the culture dish. The cells are incubated for 5 t 6 hours at 37°C/5% CO_2. After the medium is aspirated off, 5 ml of fresh medium containing 2% of foetal bovine serum are add d. After 48 h urs at 37°C/5% CO_2, th cells are rinsed with PBS, detached by trypsinization, dilut d in 10 ml of culture medium (5%

foetal bovine serum) and plated out in a dish 10 cm in diameter (the dilution may be adjusted in accordance with the efficiency of transfection). After a further incubation for 10 hours (the time for the cells to adhere), the cells are subjected to selection by adding G418 at a final concentration of 600 µg/ml Geneticin equivalent for 15 to 21 days (the medium is changed every day). The clones obtained are then rinsed with PBS, fixed in 70% ethanol, dried, stained with 1% crystal violet and then counted.

Four plasmid transfections were carried out in duplicate:

- plasmid pCDNA3 without insert
- plasmid pCDNA3/SR-p70 containing human SR-p70a cDNA
- plasmid pCDNA3/SR-p70 Mut containing SR-p70a cDNA possessing a mutation at position 293 AA (R → H) which is analogous to the mutation 273 (R → H) in the DNA-binding domain of p53
- control without plasmid.

The result is expressed as the number of clones per dish.

	Experiment 1	Experiment 2	Mean
pCDNA3	172	353	262
pCDNA3/SR-p70	13	8	10
pCDNA3/SR-p70 Mut	92	87	89
Absence of plasmid	1	3	2

The number of clones obtained by transfection with plasmid pCDNA3/SR-p70 is 25-fold less than the number of clones obtained with the control pCDNA3 and 9-fold less than the number of clones obtained with pCDNA3/SR-p70 Mut, indicating a mortality or an arrest of cell division of the cells transfected with SR-p70 cDNA. This result is not the consequence of a toxicity in vi w of the clones obtained with the mutated SR-p70 cDNA, but probably of an apoptosis as has been demonstrated for the

p53 protein (Koshland et al., Sciences, 1993, 262, 1953-1981).

EXAMPLE XIII

Biological role of the SR-p70 protein

The structural homology between the DNA-binding domain of p53 and the central region of the SR-p70 protein enables it to be inferred that SR-p70 is a transcription factor (see Figures 1 and 2). In effect, p53 (393 amino acids) consists of several functional domains. The N-terminal region (1-91 amino acids) is involved in the activation of transcription, and contains sites for interaction with different cellular and viral proteins. The central portion (amino acids 92 to 292) permits binding to the specific DNA sequences located in the promoter regions of certain genes (the majority of point mutations that inactivate p53 are localized in this region), and also possesses numerous sites for interaction with viral proteins which inhibit its activity. Finally, the last 100 amino acids of p53 are responsible for its oligomerization as well as for the regulation of the latter (Hainaut P., Current Opinion in Oncology, 1995, 7, 76-82; Prokocimer M., Blood, 1994, *84* No. 8, 2391-2411).

The sequence homology between p53 and SR-p70 is significant, in particular as regards the amino acids involved directly in the interaction with DNA, suggesting that SR-p70 binds to the p53 sites on DNA. These amino acids correspond very exactly to what are referred to as the "hot spots", amino acids frequently mutated in human tumours (SWISS PROT: SW: P53_human and Prokocimer M., Blood, 1994, *84* No. 8, 2391-2411). From this homology, it may be deduced that the SR-p70 protein exerts a control over the activity of the gen s regulated by p53, ither independently of the latter or by forming heterooligomers with it.

Consequently, like p53, the products of the SR-p70 gene must be involved in the control and regulation

of the cell cycle, causing the cycle to stop (momentarily or permanently), and the implementation of programmes such as DNA repair, differentiation or cell death. The likelihood of the existence of "p53-like" activities had been strongly felt with the demonstration in $p53^{-/-}$ mic of activities of DNA repair and cell death in response to ionizing radiations (Strasser *et al.*, Cell, 1994, *79*, 329-339). The authors of the present invention hav localized the human SR-p70 gene in the telomeric region of the short arm of chromosome 1, precisely at 1p36.2-36.3, the smallest deleted region (SRO) common to a majority of neuroblastomas and of other types of tumours (melanomas and sarcomas) (White *et al.*, PNAS, 1995, *92*, 5520-5524). This region of loss of heterozygosity (LOH) defines the locus of a tumour-suppressing gene whose loss of activity is considered to be the cause of tumour formation. It is important to recall that this region is also subject to "maternal imprinting"; the maternal allele is preferentially lost in neuroblastomas having the 1p36 deletion (without amplification of N-Myc) (Caron *et al.*, Hum. Mol. Gen., 1995, *4*, 535-539). The wide-type SR-p70 gene introduced into neuroblastoma cells and expressed therein permits the reversion of their transformation. The loss of this anti-oncogenic activity is hence associated with the development of the tumour. The 1p36 region possesses a syngeneic homology with the distal segment of the mouse chromosome 4. In this region, the *curly tail (ct)* gene (Beier *et al.*, Mammalian Genome, 1995, *6*, 269-272) involved in congenital malformations of the neural tube (NTM: *spina bifida*, anencephaly, etc). The *ct* mouse is the best animal model for studying these malformations. It is accepted that these malformations result from abnormalities of cell proliferation. Bearing in mind the nature of the SR-p70 gene and its chromosomal localization, one of the hypotheses is that SR-p70 could be the human homologue of *ct* and that, on this basis, the detection of early mutations and chromosomal abnormaliti s affecting this g n should permit, for example, as an application, th identification of persons

at risk (0.5-1% of newborn babies affected by NTM) and the implementation of preventive treatments (Neumann *et al.*, Nature Genetics, 1994, *6*, 357-362; Di Vinci *et al.*, Int. J. Cancer, 1994, *59*, 422-426; Moll *et al.*, PNAS, 1995, *92*, 4407-4411; Chen *et al.*, Development, 1995, *121*, 681-691).

EXAMPLE XIV

Allelic study of the SR-p70 gene

The GC and AT alleles are readily identified by StyI restriction of the PCR products of exon 2 (see Example XI). Hence it was possible to determine in this way, in GC/AT heterozygous individuals bearing neuroblastoma tumours, the lost SR-p70 allele (GC or AT), in spite of the presence of contaminating healthy tissue.

Surprisingly, when the same analysis is carried out on the RNA, a single allele is demonstrated independently of the presence or otherwise of a deletion and, still more surprisingly, in spite of the presence of healthy tissue. This suggests that the imprint (differential expression of the two alleles) would also exist in the contaminating tissue.

In order to verify this, the same analysis was repeated on the RNA originating from blood cells of healthy GC/AT heterozygous individuals. Only one of the two types of transcript was detected also in these cells. This result confirms the observation made on the tumour samples regarding the existence of a generalized genetic imprint for the SR-p70 gene.

The implications of this discovery are important, since it enables it to be postulated that a single sporadic mutation inactivating the active SR-p70 allele will give rise to a loss of activity, this potentially occurring in all the tissues.

The absence of precise data on the biological function of SR-p70 does not enable the consequences of this loss of SR-p70 activity for the cell to be measured.

Nevertheless, its strong homology with the p53 tumour-suppressing protein, as well as the demonstration that SR-p70 is a transcription factor capable of utilizing th $P21^{waf}$ promoter, suggests a role of this protein in th control of the cell cycle and in differentiation.

Knudson and Meadows, 1980 (New Eng. J. Med. 302: 1254-56), consider the IV-S neuroblastomas to be a collection of non-malignant cells from the neural crest carrying a mutation which interferes with their normal differentiation.

It is conceivable that the loss of SR-p70 activity, like the loss of p53 control over the cell cycle, favours the appearance of cellular abnormalities such as aneuploidy, amplification (described in the cas of neuroblastomas) and other genetic reorganizations capable of causing cell transformation (Livingstone et al., 1992, Cell 71:923-25; Yin et al. 1992, Cell 72:937-48; Cross et al. 1995, Science 267:1353-56; Fukasawa et al. 1996, Science 271:1744-47). Neuroblastomas might hence arise originally from a temporary or permanent loss of activity of SR-p70, thereby favouring the occurrenc of oncogenic events and hence tumour progression.

In the case of the 1p36 constitutional deletion described by Biegel et al., 1993 (Am. J. Hum. Genet. 52:176-82), IV-S neuroblastoma does indeed occur and the gene affected is NBS-1 (SR-p70).

In conclusion, what is described for neuroblastomas might also apply to other types of tumours, in particular those associated with reorganization of the end of the short arm of chromosome 1 (Report 2 international workshop on human chr 1 mapping 1995, Cytogenetics and Cell Genet. 72:113-154). From a therapeutic standpoint, the involvement of SR-p70 in th occurrence of tumours should lead to the avoidance of the use of mutag nic agents in chemotherapy, bearing in mind th risks of cell transformation by these products, and to the use, in preference to thes products, of non-mutagenic substanc s which stimulate differentiation.

Moreov r, the frequency of occurrence f the GC

and AT alleles is as follows: in the population, Frequency(AT)=0.15, and on a sample of 25 (neuroblastoma) patients, F(AT)=0.30. These statistics indicate that the AT allele could be a predisposing factor.